"A Generalized Mechanism for Unicode Metadata"

ABSTRACT OF THE DISCLOSURE

[0122]   A extendable method for including display rendering metadata within

Unicode character streams.   Metadata is distinct from character data, even though it

is embedded in the Unicode character stream using tag mechanism. The method

allows for an unlimited number of tag identifiers. Legacy Unicode methods such as

Bidi, Normalization, and Line Breaking, can be recast using the invention in a more

manageable context according to the metadata framework, thereby allowing the

methods to be detectable, reversible as well as convertible. The traditional Unicode

Control Layer is eliminated because the syntax of controls are captured universally by

the new Metadata Layer, irrespective of whether the control relates to presentation or

pcontent. By replacing the indistinct boundary separating characters and control with

a well defined division, applications that rely on Unicode are easier to develop and to

maintain.